

Announcements

- Assignment drop-in this Thursday, 2pm (questions 4, 9, 10, 11, 12). Good place for hints, tips etc.
- First computer practical this Friday, 1pm, Herschel PC cluster
- ReCap problems...?
- Next week: 2 lectures and a problems class (case study 1)

Example 2.4

Table 2.3 shows some data on the times between serious earthquakes.

An earthquake is included if its magnitude is at least 7.5 on the Richter scale or if over 1000 people were killed.

Recording starts on 16 December 1902 (4500 killed in Turkistan).

The table includes data on 21 earthquakes, that is, 20 “waiting times” between earthquakes.

840	157	145	44	33	121	150	280	434	736
584	887	263	1901	695	294	562	721	76	710

Example 2.4

It is believed that earthquakes happen in a random haphazard kind of way and that times between earthquakes can be described by an exponential distribution.

Data over a much longer period suggest that this exponential assumption is plausible.

Therefore, we will assume that these data are a random sample from an exponential distribution with rate θ (and mean $1/\theta$). The parameter θ describes the rate at which earthquakes occur.

Example 2.4

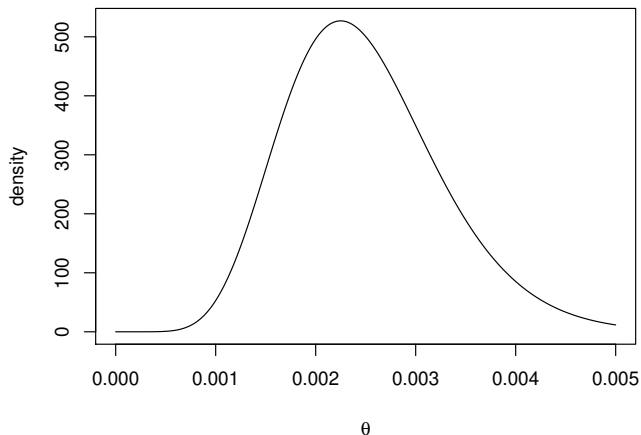
An expert on earthquakes has prior beliefs about the rate of earthquakes, θ , described by a $Ga(10, 4000)$ distribution (see Example 3.1, Chapter 3), which has density density

$$\pi(\theta) = \frac{4000^{10} \theta^9 e^{-4000\theta}}{\Gamma(10)}, \quad \theta > 0, \quad (2.13)$$

and mean $E(\theta) = 0.0025$.

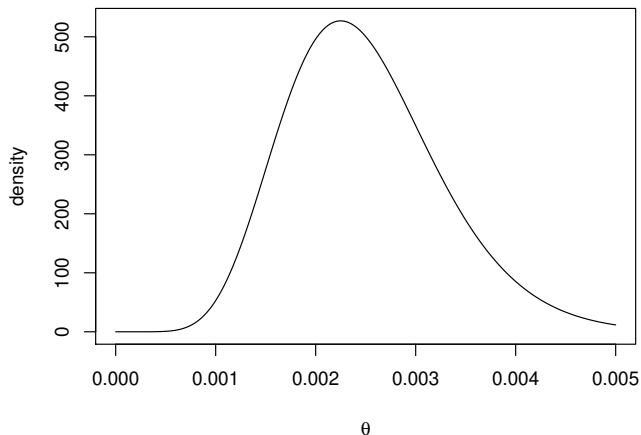
Example 2.4

A plot of this prior distribution can be found in Figure 2.8. As you might expect, the expert believes that only very small values of θ are likely, though larger values are not ruled out!



Example 2.4

A plot of this prior distribution can be found in Figure 2.8. As you might expect, the expert believes that only very small values of θ are likely, though larger values are not ruled out!



Solution to Example 2.4 (1/2)

The data are observations on $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, 20$ (independent).

Therefore, the likelihood function for θ is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^{20} \theta e^{-\theta x_i}, \quad \theta > 0 \\ &= \theta^{20} \exp\left(-\theta \sum_{i=1}^{20} x_i\right), \quad \theta > 0 \\ &= \theta^{20} e^{-9633\theta}, \quad \theta > 0. \end{aligned} \tag{2.14}$$

Solution to Example 2.4 (1/2)

The data are observations on $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, 20$ (independent).

Therefore, the likelihood function for θ is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^{20} \theta e^{-\theta x_i}, \quad \theta > 0 \\ &= \theta^{20} \exp\left(-\theta \sum_{i=1}^{20} x_i\right), \quad \theta > 0 \\ &= \theta^{20} e^{-9633\theta}, \quad \theta > 0. \end{aligned} \tag{2.14}$$

Solution to Example 2.4 (1/2)

The data are observations on $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, 20$ (independent).

Therefore, the likelihood function for θ is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^{20} \theta e^{-\theta x_i}, \quad \theta > 0 \\ &= \theta^{20} \exp\left(-\theta \sum_{i=1}^{20} x_i\right), \quad \theta > 0 \\ &= \theta^{20} e^{-9633\theta}, \quad \theta > 0. \end{aligned} \tag{2.14}$$

Solution to Example 2.4 (2/2)

We now apply Bayes Theorem to combine the expert opinion with the observed data. The posterior density function is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \frac{4000^{10} \theta^9 e^{-4000\theta}}{\Gamma(10)} \times \theta^{20} e^{-9633\theta}, \quad \theta > 0 \\ &= k \theta^{30-1} e^{-13633\theta}, \quad \theta > 0.\end{aligned}\tag{2.15}$$

The only continuous distribution which takes the form $k\theta^{g-1}e^{-h\theta}$, $\theta > 0$ is the $Ga(g, h)$ distribution.

Therefore, the posterior distribution must be $\theta|\mathbf{x} \sim Ga(30, 13633)$.

Solution to Example 2.4 (2/2)

We now apply Bayes Theorem to combine the expert opinion with the observed data. The posterior density function is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \frac{4000^{10} \theta^9 e^{-4000\theta}}{\Gamma(10)} \times \theta^{20} e^{-9633\theta}, \quad \theta > 0 \\ &= k \theta^{30-1} e^{-13633\theta}, \quad \theta > 0.\end{aligned}\tag{2.15}$$

The only continuous distribution which takes the form $k\theta^{g-1}e^{-h\theta}$, $\theta > 0$ is the $Ga(g, h)$ distribution.

Therefore, the posterior distribution must be $\theta|\mathbf{x} \sim Ga(30, 13633)$.

Solution to Example 2.4 (2/2)

We now apply Bayes Theorem to combine the expert opinion with the observed data. The posterior density function is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \frac{4000^{10} \theta^9 e^{-4000\theta}}{\Gamma(10)} \times \theta^{20} e^{-9633\theta}, \quad \theta > 0 \\ &= k \theta^{30-1} e^{-13633\theta}, \quad \theta > 0.\end{aligned}\tag{2.15}$$

The only continuous distribution which takes the form $k\theta^{g-1}e^{-h\theta}$, $\theta > 0$ is the $Ga(g, h)$ distribution.

Therefore, the posterior distribution must be $\theta|\mathbf{x} \sim Ga(30, 13633)$.

Solution to Example 2.4 (2/2)

We now apply Bayes Theorem to combine the expert opinion with the observed data. The posterior density function is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \frac{4000^{10} \theta^9 e^{-4000\theta}}{\Gamma(10)} \times \theta^{20} e^{-9633\theta}, \quad \theta > 0 \\ &= k \theta^{30-1} e^{-13633\theta}, \quad \theta > 0.\end{aligned}\tag{2.15}$$

The only continuous distribution which takes the form $k\theta^{g-1}e^{-h\theta}$, $\theta > 0$ is the $Ga(g, h)$ distribution.

Therefore, the posterior distribution must be $\theta|\mathbf{x} \sim Ga(30, 13633)$.

Example 2.4: Summary

The data have updated our beliefs about θ from a $Ga(10, 4000)$ distribution to a $Ga(30, 13633)$ distribution.

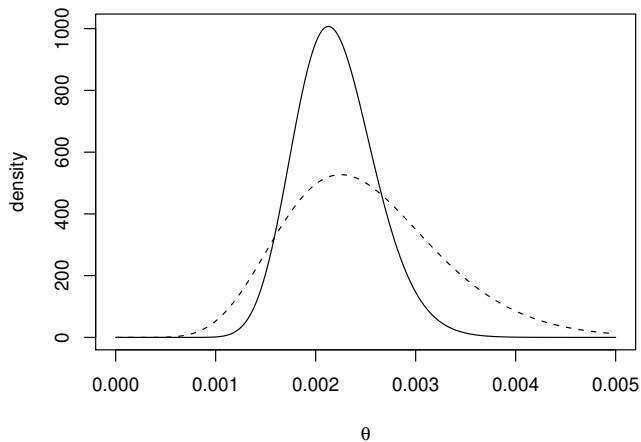
Plots of these distributions are given in Figure 2.9, and Table 2.4 gives a summary of the main changes induced by incorporating the data.

Example 2.4: Summary

The data have updated our beliefs about θ from a $Ga(10, 4000)$ distribution to a $Ga(30, 13633)$ distribution.

Plots of these distributions are given in Figure 2.9, and Table 2.4 gives a summary of the main changes induced by incorporating the data.

Example 2.4: Summary



Example 2.4: Summary

	Prior (2.13)	Likelihood (2.14)	Posterior (2.15)
$Mode(\theta)$	0.00225	0.00208	0.00213
$E(\theta)$	0.00250	—	0.00220
$SD(\theta)$	0.00079	—	0.00040

Notice that, as the mode of the likelihood function is close to that of the prior distribution, the information in the data is consistent with that in the prior distribution.

This results in a reduction in variability from the prior to the posterior distributions.

The similarity between the prior beliefs and the data has reduced the uncertainty we have about the likely earthquake rate θ .

Example 2.4: Summary

	Prior (2.13)	Likelihood (2.14)	Posterior (2.15)
$Mode(\theta)$	0.00225	0.00208	0.00213
$E(\theta)$	0.00250	—	0.00220
$SD(\theta)$	0.00079	—	0.00040

Notice that, as the mode of the likelihood function is close to that of the prior distribution, the information in the data is consistent with that in the prior distribution.

This results in a reduction in variability from the prior to the posterior distributions.

The similarity between the prior beliefs and the data has reduced the uncertainty we have about the likely earthquake rate θ .

Example 2.4: Summary

	Prior (2.13)	Likelihood (2.14)	Posterior (2.15)
$Mode(\theta)$	0.00225	0.00208	0.00213
$E(\theta)$	0.00250	—	0.00220
$SD(\theta)$	0.00079	—	0.00040

Notice that, as the mode of the likelihood function is close to that of the prior distribution, the information in the data is consistent with that in the prior distribution.

This results in a reduction in variability from the prior to the posterior distributions.

The similarity between the prior beliefs and the data has reduced the uncertainty we have about the likely earthquake rate θ .

Example 2.4: Summary

	Prior (2.13)	Likelihood (2.14)	Posterior (2.15)
$Mode(\theta)$	0.00225	0.00208	0.00213
$E(\theta)$	0.00250	—	0.00220
$SD(\theta)$	0.00079	—	0.00040

Notice that, as the mode of the likelihood function is close to that of the prior distribution, the information in the data is consistent with that in the prior distribution.

This results in a reduction in variability from the prior to the posterior distributions.

The similarity between the prior beliefs and the data has reduced the uncertainty we have about the likely earthquake rate θ .

Example 2.5

We now consider the general case of the problem discussed in Example 2.4.

Suppose $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, n$ (independent) and our prior beliefs about θ are summarised by a $\text{Ga}(g, h)$ distribution (with g and h known), with density

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0. \quad (2.16)$$

Determine the posterior distribution for θ .

Solution to Example 2.5 (1/2)

The likelihood function for θ is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i}, \quad \theta > 0 \\&= \theta^n e^{-\theta x_1 - \theta x_2 - \dots - \theta x_n} \\&= \theta^n e^{-\theta \sum x} \\&= \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.\end{aligned}\tag{2.17}$$

Solution to Example 2.5 (2/2)

We now apply Bayes Theorem. The posterior density function is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.$$

$$\pi(\theta|\mathbf{x}) = k\theta^{g+n-1} e^{-(h+n\bar{x})\theta}, \quad \theta > 0, \quad (2.18)$$

where k is a constant that does not depend on θ .

Therefore, the posterior distribution takes the form $k\theta^{g-1} e^{-h\theta}$, $\theta > 0$ and so must be a gamma distribution.

Thus we have $\theta|\mathbf{x} \sim Ga(g+n, h+n\bar{x})$.

Solution to Example 2.5 (2/2)

We now apply Bayes Theorem. The posterior density function is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.$$

$$\pi(\theta|\mathbf{x}) = k\theta^{g+n-1} e^{-(h+n\bar{x})\theta}, \quad \theta > 0, \quad (2.18)$$

where k is a constant that does not depend on θ .

Therefore, the posterior distribution takes the form $k\theta^{g-1} e^{-h\theta}$, $\theta > 0$ and so must be a gamma distribution.

Thus we have $\theta|\mathbf{x} \sim Ga(g+n, h+n\bar{x})$.

Solution to Example 2.5 (2/2)

We now apply Bayes Theorem. The posterior density function is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.$$

$$\pi(\theta|\mathbf{x}) = k\theta^{g+n-1} e^{-(h+n\bar{x})\theta}, \quad \theta > 0, \quad (2.18)$$

where k is a constant that does not depend on θ .

Therefore, the posterior distribution takes the form $k\theta^{g-1} e^{-h\theta}$, $\theta > 0$ and so must be a gamma distribution.

Thus we have $\theta|\mathbf{x} \sim Ga(g+n, h+n\bar{x})$.

Solution to Example 2.5 (2/2)

We now apply Bayes Theorem. The posterior density function is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.$$

$$\pi(\theta|\mathbf{x}) = k\theta^{g+n-1} e^{-(h+n\bar{x})\theta}, \quad \theta > 0, \quad (2.18)$$

where k is a constant that does not depend on θ .

Therefore, the posterior distribution takes the form $k\theta^{g-1} e^{-h\theta}$, $\theta > 0$ and so must be a gamma distribution.

Thus we have $\theta|\mathbf{x} \sim Ga(g+n, h+n\bar{x})$.

Solution to Example 2.5 (2/2)

We now apply Bayes Theorem. The posterior density function is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \theta^n e^{-n\bar{x}\theta}, \quad \theta > 0.$$

$$\pi(\theta|\mathbf{x}) = k\theta^{g+n-1} e^{-(h+n\bar{x})\theta}, \quad \theta > 0, \quad (2.18)$$

where k is a constant that does not depend on θ .

Therefore, the posterior distribution takes the form $k\theta^{g-1} e^{-h\theta}$, $\theta > 0$ and so must be a gamma distribution.

Thus we have $\theta|\mathbf{x} \sim Ga(g+n, h+n\bar{x})$.

Example 2.5: Summary

If

- $X_i \sim \text{Exp}(\theta)$ and
- $\theta \sim \text{Ga}(g, h)$, then
- $\theta | \mathbf{x} \sim \text{Ga}(g + n, h + n\bar{x})$

The changes in our beliefs about θ are summarised in Table 2.5, taking $g \geq 1$.

	Prior (2.16)	Likelihood (2.17)	Posterior (2.18)
$\text{Mode}(\theta)$	$(g - 1)/h$	$1/\bar{x}$	$(g + n - 1)/(h + n\bar{x})$
$E(\theta)$	g/h	—	$(g + n)/(h + n\bar{x})$
$SD(\theta)$	$\sqrt{g}/h$	—	$\sqrt{g + n}/(h + n\bar{x})$

Example 2.5: Summary

If

- $X_i \sim \text{Exp}(\theta)$ and
- $\theta \sim \text{Ga}(g, h)$, then
- $\theta | \mathbf{x} \sim \text{Ga}(g + n, h + n\bar{x})$

The changes in our beliefs about θ are summarised in Table 2.5, taking $g \geq 1$.

	Prior (2.16)	Likelihood (2.17)	Posterior (2.18)
$\text{Mode}(\theta)$	$(g - 1)/h$	$1/\bar{x}$	$(g + n - 1)/(h + n\bar{x})$
$E(\theta)$	g/h	—	$(g + n)/(h + n\bar{x})$
$SD(\theta)$	$\sqrt{g}/h$	—	$\sqrt{g + n}/(h + n\bar{x})$

Example 2.5: Summary

If

- $X_i \sim \text{Exp}(\theta)$ and
- $\theta \sim \text{Ga}(g, h)$, then
- $\theta | \mathbf{x} \sim \text{Ga}(g + n, h + n\bar{x})$

The changes in our beliefs about θ are summarised in Table 2.5, taking $g \geq 1$.

	Prior (2.16)	Likelihood (2.17)	Posterior (2.18)
$\text{Mode}(\theta)$	$(g - 1)/h$	$1/\bar{x}$	$(g + n - 1)/(h + n\bar{x})$
$E(\theta)$	g/h	—	$(g + n)/(h + n\bar{x})$
$SD(\theta)$	$\sqrt{g}/h$	—	$\sqrt{g + n}/(h + n\bar{x})$

Example 2.5: Summary

If

- $X_i \sim \text{Exp}(\theta)$ and
- $\theta \sim \text{Ga}(g, h)$, then
- $\theta | \mathbf{x} \sim \text{Ga}(g + n, h + n\bar{x})$

The changes in our beliefs about θ are summarised in Table 2.5, taking $g \geq 1$.

	Prior (2.16)	Likelihood (2.17)	Posterior (2.18)
$\text{Mode}(\theta)$	$(g - 1)/h$	$1/\bar{x}$	$(g + n - 1)/(h + n\bar{x})$
$E(\theta)$	g/h	—	$(g + n)/(h + n\bar{x})$
$SD(\theta)$	$\sqrt{g}/h$	—	$\sqrt{g + n}/(h + n\bar{x})$

Chapter 2 so far...

Model	Prior	Posterior
Binomial(k, θ)	$\theta \sim U(0, 1)$ $\theta \sim \text{Beta}(a, b)$	Beta Beta
Exponential(θ)	$\theta \sim \text{Ga}(g, h)$ $\theta \sim \text{exp}(g)$	Gamma ??
Normal(μ, σ^2)	$\mu \sim N(b, 1/d)$	Normal

Chapter 2 so far...

Model	Prior	Posterior
Binomial(k, θ)	$\theta \sim U(0, 1)$ $\theta \sim \text{Beta}(a, b)$	Beta Beta
Exponential(θ)	$\theta \sim \text{Ga}(g, h)$ $\theta \sim \exp(g)$	Gamma ??
Normal(μ, σ^2)	$\mu \sim N(b, 1/d)$	Normal

Chapter 2 so far...

Model	Prior	Posterior
Binomial(k, θ)	$\theta \sim U(0, 1)$ $\theta \sim \text{Beta}(a, b)$	Beta Beta
Exponential(θ)	$\theta \sim \text{Ga}(g, h)$ $\theta \sim \text{exp}(g)$	Gamma ??
Normal(μ, σ^2)	$\mu \sim N(b, 1/d)$	Normal

Chapter 2 so far...

Model	Prior	Posterior
Binomial(k, θ)	$\theta \sim U(0, 1)$ $\theta \sim \text{Beta}(a, b)$	Beta Beta
Exponential(θ)	$\theta \sim \text{Ga}(g, h)$ $\theta \sim \text{exp}(g)$	Gamma ??
Normal(μ, σ^2)	$\mu \sim N(b, 1/d)$	Normal

Definition (Conjugacy)

Suppose that data $\mathbf{x}$ are to be observed with distribution $f(\mathbf{x}|\theta)$.

A family $\mathfrak{F}$ of prior distributions for θ is said to be **conjugate** to $f(\mathbf{x}|\theta)$ if for every prior distribution $\pi(\theta) \in \mathfrak{F}$, the posterior distribution $\pi(\theta|\mathbf{x})$ is also in $\mathfrak{F}$.

Example 2.6

Suppose we have a random sample from a Normal distribution.

In Bayesian statistics, when dealing with the Normal distribution, the mathematics is more straightforward if we work with the **precision** ($= 1/\text{variance}$) of the distribution rather than the variance itself.

So we will assume that this population has unknown mean μ but known precision τ : $X_i|\mu \sim N(\mu, 1/\tau)$, $i = 1, 2, \dots, n$ (independent), where τ is known.

Suppose our prior beliefs about μ can be summarised by a $N(b, 1/d)$ distribution, with probability density function

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \quad (2.19)$$

Determine the posterior distribution for μ .

Example 2.6

Suppose we have a random sample from a Normal distribution.

In Bayesian statistics, when dealing with the Normal distribution, the mathematics is more straightforward if we work with the **precision** ($= 1/\text{variance}$) of the distribution rather than the variance itself.

So we will assume that this population has unknown mean μ but known precision τ : $X_i|\mu \sim N(\mu, 1/\tau)$, $i = 1, 2, \dots, n$ (independent), where τ is known.

Suppose our prior beliefs about μ can be summarised by a $N(b, 1/d)$ distribution, with probability density function

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \quad (2.19)$$

Determine the posterior distribution for μ .

Example 2.6

Suppose we have a random sample from a Normal distribution.

In Bayesian statistics, when dealing with the Normal distribution, the mathematics is more straightforward if we work with the **precision** ($= 1/\text{variance}$) of the distribution rather than the variance itself.

So we will assume that this population has unknown mean μ but known precision τ : $X_i|\mu \sim N(\mu, 1/\tau)$, $i = 1, 2, \dots, n$ (independent), where τ is known.

Suppose our prior beliefs about μ can be summarised by a $N(b, 1/d)$ distribution, with probability density function

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \quad (2.19)$$

Determine the posterior distribution for μ .

Example 2.6

Suppose we have a random sample from a Normal distribution.

In Bayesian statistics, when dealing with the Normal distribution, the mathematics is more straightforward if we work with the **precision** ($= 1/\text{variance}$) of the distribution rather than the variance itself.

So we will assume that this population has unknown mean μ but known precision τ : $X_i|\mu \sim N(\mu, 1/\tau)$, $i = 1, 2, \dots, n$ (independent), where τ is known.

Suppose our prior beliefs about μ can be summarised by a $N(b, 1/d)$ distribution, with probability density function

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \quad (2.19)$$

Determine the posterior distribution for μ .

Solution to Example 2.6 (1/5)

The likelihood function for μ is

$$\begin{aligned}f(\mathbf{x}|\mu) &= \prod_{i=1}^n \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n ([x_i - \bar{x}] + [\bar{x} - \mu])^2\right\}\end{aligned}$$

Solution to Example 2.6 (1/5)

The likelihood function for μ is

$$\begin{aligned}f(\mathbf{x}|\mu) &= \prod_{i=1}^n \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n ([x_i - \bar{x}] + [\bar{x} - \mu])^2\right\}\end{aligned}$$

Solution to Example 2.6 (1/5)

The likelihood function for μ is

$$\begin{aligned}f(\mathbf{x}|\mu) &= \prod_{i=1}^n \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n ([x_i - \bar{x}] + [\bar{x} - \mu])^2\right\}\end{aligned}$$

Solution to Example 2.6 (1/5)

The likelihood function for μ is

$$\begin{aligned}f(\mathbf{x}|\mu) &= \prod_{i=1}^n \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2} \sum_{i=1}^n ([x_i - \bar{x}] + [\bar{x} - \mu])^2\right\}\end{aligned}$$

Solution to Example 2.6 (2/5)

Giving

$$\begin{aligned} f(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right. \right. \\ &\quad \left. \left. + 2(\bar{x} - \mu) \sum_{i=1}^n (x_i - \bar{x}) + \sum_{i=1}^n (\bar{x} - \mu)^2 \right] \right\} \\ &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \right] \right\}, \end{aligned}$$

Let

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2;$$

Then

$$f(\mathbf{x}|\mu) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{n\tau}{2} [s^2 + (\bar{x} - \mu)^2] \right\}. \quad (2.20)$$

Solution to Example 2.6 (2/5)

Giving

$$\begin{aligned}f(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right. \right. \\&\quad \left. \left. + 2(\bar{x} - \mu) \sum_{i=1}^n (x_i - \bar{x}) + \sum_{i=1}^n (\bar{x} - \mu)^2 \right] \right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \right] \right\},\end{aligned}$$

Let

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2;$$

Then

$$f(\mathbf{x}|\mu) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{n\tau}{2} \left[s^2 + (\bar{x} - \mu)^2 \right] \right\}. \quad (2.20)$$

Solution to Example 2.6 (2/5)

Giving

$$\begin{aligned} f(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right. \right. \\ &\quad \left. \left. + 2(\bar{x} - \mu) \sum_{i=1}^n (x_i - \bar{x}) + \sum_{i=1}^n (\bar{x} - \mu)^2 \right] \right\} \\ &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \right] \right\}, \end{aligned}$$

Let

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2;$$

Then

$$f(\mathbf{x}|\mu) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{n\tau}{2} \left[s^2 + (\bar{x} - \mu)^2 \right] \right\}. \quad (2.20)$$

Solution to Example 2.6 (2/5)

Giving

$$\begin{aligned} f(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right. \right. \\ &\quad \left. \left. + 2(\bar{x} - \mu) \sum_{i=1}^n (x_i - \bar{x}) + \sum_{i=1}^n (\bar{x} - \mu)^2 \right] \right\} \\ &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{\tau}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \right] \right\}, \end{aligned}$$

Let

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2;$$

Then

$$f(\mathbf{x}|\mu) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left\{ -\frac{n\tau}{2} \left[s^2 + (\bar{x} - \mu)^2 \right] \right\}. \quad (2.20)$$

Solution to Example 2.6 (3/5)

Applying Bayes Theorem, the posterior density function is

$$\pi(\mu|\mathbf{x}) \propto \pi(\mu)f(\mathbf{x}|\mu)$$

$$\begin{aligned} &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\} \end{aligned}$$

$$= k_1 \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\}$$

$$= k_1 \exp\left\{-\frac{1}{2}\left[d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2\right]\right\}$$

where k_1 is a constant that does not depend on μ .

Solution to Example 2.6 (3/5)

Applying Bayes Theorem, the posterior density function is

$$\pi(\mu|\mathbf{x}) \propto \pi(\mu)f(\mathbf{x}|\mu)$$

$$\begin{aligned} &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\} \end{aligned}$$

$$= k_1 \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\}$$

$$= k_1 \exp\left\{-\frac{1}{2}\left[d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2\right]\right\}$$

where k_1 is a constant that does not depend on μ .

Solution to Example 2.6 (3/5)

Applying Bayes Theorem, the posterior density function is

$$\pi(\mu|\mathbf{x}) \propto \pi(\mu)f(\mathbf{x}|\mu)$$

$$\begin{aligned} &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\} \end{aligned}$$

$$= k_1 \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\}$$

$$= k_1 \exp\left\{-\frac{1}{2}\left[d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2\right]\right\}$$

where k_1 is a constant that does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (4/5)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 &+ n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 + n\tau\mu^2 - 2db\mu - 2n\tau\bar{x}\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau)\mu^2 - 2(db + n\tau\bar{x})\mu + db^2 + n\tau\bar{x}^2 \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + c \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Solution to Example 2.6 (5/5)

Let

$$B = \frac{db + n\tau\bar{x}}{d + n\tau} \quad \text{and} \quad D = d + n\tau. \quad (2.21)$$

Then

$$\begin{aligned} \pi(\mu|\mathbf{x}) &= k_1 \exp \left\{ -\frac{D}{2}(\mu - B)^2 - \frac{c}{2} \right\} \\ &= k \exp \left\{ -\frac{D}{2}(\mu - B)^2 \right\}, \end{aligned} \quad (2.22)$$

where k is a constant that does not depend on μ .

Therefore, the posterior distribution takes the form $k \exp\{-D(\mu - B)^2/2\}$, $-\infty < \mu < \infty$ and so must be a normal distribution: we have $\mu|\mathbf{x} \sim N(B, 1/D)$.

Solution to Example 2.6 (5/5)

Let

$$B = \frac{db + n\tau\bar{x}}{d + n\tau} \quad \text{and} \quad D = d + n\tau. \quad (2.21)$$

Then

$$\begin{aligned} \pi(\mu|\mathbf{x}) &= k_1 \exp \left\{ -\frac{D}{2}(\mu - B)^2 - \frac{c}{2} \right\} \\ &= k \exp \left\{ -\frac{D}{2}(\mu - B)^2 \right\}, \end{aligned} \quad (2.22)$$

where k is a constant that does not depend on μ .

Therefore, the posterior distribution takes the form $k \exp\{-D(\mu - B)^2/2\}$, $-\infty < \mu < \infty$ and so must be a normal distribution: we have $\mu|\mathbf{x} \sim N(B, 1/D)$.

Solution to Example 2.6 (5/5)

Let

$$B = \frac{db + n\tau\bar{x}}{d + n\tau} \quad \text{and} \quad D = d + n\tau. \quad (2.21)$$

Then

$$\begin{aligned} \pi(\mu|\mathbf{x}) &= k_1 \exp \left\{ -\frac{D}{2}(\mu - B)^2 - \frac{c}{2} \right\} \\ &= k \exp \left\{ -\frac{D}{2}(\mu - B)^2 \right\}, \end{aligned} \quad (2.22)$$

where k is a constant that does not depend on μ .

Therefore, the posterior distribution takes the form $k \exp\{-D(\mu - B)^2/2\}$, $-\infty < \mu < \infty$ and so must be a normal distribution: we have $\mu|\mathbf{x} \sim N(B, 1/D)$.

Example 2.6: Summary

If we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known) and our prior beliefs about μ follow a $N(b, 1/d)$ distribution then, after incorporating the data, our (posterior) beliefs about μ follow a $N(B, 1/D)$ distribution.

Example 2.6: Summary

If we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known) and our prior beliefs about μ follow a $N(b, 1/d)$ distribution then, after incorporating the data, our (posterior) beliefs about μ follow a $N(B, 1/D)$ distribution.

Example 2.6: Summary

The changes in our beliefs about μ are summarised in Table 2.6.

	Prior (2.19)	Likelihood (2.20)	Posterior (2.21)
$Mode(\mu)$	b	$\bar{x}$	$(db + n\tau\bar{x})/(d + n\tau)$
$E(\mu)$	b	—	$(db + n\tau\bar{x})/(d + n\tau)$
$Precision(\mu)$	d	—	$d + n\tau$

Example 2.6: Summary

The changes in our beliefs about μ are summarised in Table 2.6.

	Prior (2.19)	Likelihood (2.20)	Posterior (2.21)
$Mode(\mu)$	b	$\bar{x}$	$(db + n\tau\bar{x})/(d + n\tau)$
$E(\mu)$	b	—	$(db + n\tau\bar{x})/(d + n\tau)$
$Precision(\mu)$	d	—	$d + n\tau$

Example 2.6: Summary

Notice that the way prior information and observed data combine is through the parameters of the normal distribution:

$$b \rightarrow \frac{db + n\tau\bar{x}}{d + n\tau} \quad \text{and} \quad d \rightarrow d + n\tau.$$

Example 2.6: Summary

Notice that the way prior information and observed data combine is through the parameters of the normal distribution:

$$b \rightarrow \frac{db + n_{\tau}\bar{x}}{d + n_{\tau}} \quad \text{and} \quad d \rightarrow d + n_{\tau}.$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

This equation for the posterior mean, which can be rewritten as

$$E(\mu|\mathbf{x}) = \alpha E(\mu) + (1 - \alpha)\bar{x},$$

arises in other models and is known as the **Bayes linear rule** (see Problems Sheet 2, question 3).

Example 2.6: Summary

Notice that the posterior mean is greater than the prior mean if and only if the likelihood mode (sample mean) is greater than the prior mean, that is

$$E(\mu|\mathbf{x}) > E(\mu) \quad \Longleftrightarrow \quad \text{Mode}[f(\mathbf{x}|\mu)] > E(\mu).$$

Also, the standard deviation of the posterior distribution is smaller than that of the prior distribution.

Example 2.6: Summary

Notice that the posterior mean is greater than the prior mean if and only if the likelihood mode (sample mean) is greater than the prior mean, that is

$$E(\mu|\mathbf{x}) > E(\mu) \quad \Longleftrightarrow \quad \text{Mode}[f(\mathbf{x}|\mu)] > E(\mu).$$

Also, the standard deviation of the posterior distribution is smaller than that of the prior distribution.

Example 2.7

The ages of *Ennerdale granophyre* rocks can be determined using the relative proportions of rubidium–87 and strontium–87 in the rock.

An expert in the field suggests that the ages of such rocks (in millions of years) $X|\mu \sim N(\mu, 8^2)$ and that a prior distribution $\mu \sim N(370, 20^2)$ is appropriate.

A rock is found whose chemical analysis yields $x = 421$. What is the posterior distribution for μ and what is the probability that the rock will be older than 400 million years?

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

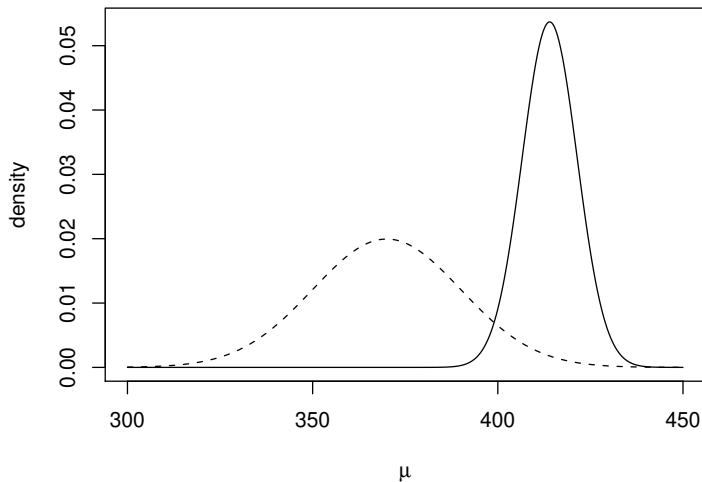
Example 2.7: Summary

This highlights the benefit of taking the chemical measurements.

Note that the large difference between these probabilities is not necessarily due to the expert's prior distribution being inaccurate, *per se*.

It is probably due to the large prior uncertainty about rock ages, as shown in Figure 2.10.

Example 2.7: Summary



We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

Posterior distributions and sufficient statistics

We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

Therefore, our (posterior) beliefs about θ having observed the full data $\mathbf{x}$ are the same as if we had observed only the sufficient statistic T .

This is what we would expect if all the information about θ in the data were contained in the sufficient statistic.

Therefore, our (posterior) beliefs about θ having observed the full data $\mathbf{x}$ are the same as if we had observed only the sufficient statistic T .

This is what we would expect if all the information about θ in the data were contained in the sufficient statistic.

Example 2.8

Suppose we have a random sample from an exponential distribution with a gamma prior distribution, that is,
 $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, n$ (independent) and $\theta \sim \text{Ga}(g, h)$.

Determine a sufficient statistic T for θ and verify that
 $\pi(\theta|\mathbf{x}) = \pi(\theta|t)$.

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned}\pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0\end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|\mathbf{x}$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned} \pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0 \end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|\mathbf{x}$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned}\pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0\end{aligned}$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|\mathbf{x}$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|\mathbf{x}$.

Example 2.9

Suppose we have a random sample from a normal distribution with known variance and a normal prior distribution for the mean parameter, that is, $X_i|\mu \sim N(\mu, 1/\tau)$, $i = 1, 2, \dots, n$ (independent) and $\mu \sim N(b, 1/d)$.

Determine a sufficient statistic T for μ and verify that $\pi(\mu|\mathbf{x}) = \pi(\mu|t)$.

Solution to Example 2.9 (1/3)

Recall from Equation (2.20) that

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2} \left[s^2 + (\bar{x} - \mu)^2\right]\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau s^2}{2}\right\} \times \exp\left\{-\frac{n\tau}{2}(\bar{x} - \mu)^2\right\} \\&= h(\mathbf{x}) g(\bar{x}, \mu)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \bar{X}$ is sufficient for μ .

Solution to Example 2.9 (1/3)

Recall from Equation (2.20) that

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2} \left[s^2 + (\bar{x} - \mu)^2\right]\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau s^2}{2}\right\} \times \exp\left\{-\frac{n\tau}{2}(\bar{x} - \mu)^2\right\} \\&= h(\mathbf{x}) g(\bar{x}, \mu)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \bar{X}$ is sufficient for μ .

Solution to Example 2.9 (1/3)

Recall from Equation (2.20) that

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\mu) &= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\} \\&= \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau s^2}{2}\right\} \times \exp\left\{-\frac{n\tau}{2}(\bar{x} - \mu)^2\right\} \\&= h(\mathbf{x}) g(\bar{x}, \mu)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \bar{X}$ is sufficient for μ .

Solution to Example 2.9 (2/3)

Now $T|\mu \sim N(\mu, 1/(n\tau))$ (by the Central Limit Theorem) and so

$$f_T(t|\mu) = \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\}.$$

Also

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\}.$$

Solution to Example 2.9 (2/3)

Now $T|\mu \sim N(\mu, 1/(n\tau))$ (by the Central Limit Theorem) and so

$$f_T(t|\mu) = \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\}.$$

Also

$$\pi(\mu) = \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\}.$$

Solution to Example 2.9 (3/3)

Therefore, by Bayes' Theorem,

$$\pi(\mu|t) \propto \pi(\mu)f(t|\mu)$$

$$\begin{aligned} &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\} \\ &\propto \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}(t - \mu)^2\right\} \\ &\quad \vdots \\ &\propto \exp\left\{-\frac{D}{2}(\mu - B)^2\right\} \end{aligned}$$

where B and D are as in Equation (2.21), with t replacing $\bar{x}$; that is, $\mu|t \sim N(B, 1/D)$, the same distribution as $\mu|\mathbf{x}$.

Solution to Example 2.9 (3/3)

Therefore, by Bayes' Theorem,

$$\pi(\mu|t) \propto \pi(\mu)f(t|\mu)$$

$$\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ \times \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\}$$

$$\propto \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}(t - \mu)^2\right\}$$

$\vdots$

$$\propto \exp\left\{-\frac{D}{2}(\mu - B)^2\right\}$$

where B and D are as in Equation (2.21), with t replacing $\bar{x}$;
that is, $\mu|t \sim N(B, 1/D)$, the same distribution as $\mu|\mathbf{x}$.

Solution to Example 2.9 (3/3)

Therefore, by Bayes' Theorem,

$$\pi(\mu|t) \propto \pi(\mu)f(t|\mu)$$

$$\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ \times \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\}$$

$$\propto \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}(t - \mu)^2\right\}$$

$$\vdots$$

$$\propto \exp\left\{-\frac{D}{2}(\mu - B)^2\right\}$$

where B and D are as in Equation (2.21), with t replacing $\bar{x}$;
that is, $\mu|t \sim N(B, 1/D)$, the same distribution as $\mu|\mathbf{x}$.

Solution to Example 2.9 (3/3)

Therefore, by Bayes' Theorem,

$$\pi(\mu|t) \propto \pi(\mu)f(t|\mu)$$

$$\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ \times \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\}$$

$$\propto \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}(t - \mu)^2\right\}$$

$$\vdots$$

$$\propto \exp\left\{-\frac{D}{2}(\mu - B)^2\right\}$$

where B and D are as in Equation (2.21), with t replacing $\bar{x}$; that is, $\mu|t \sim N(B, 1/D)$, the same distribution as $\mu|\mathbf{x}$.

Solution to Example 2.9 (3/3)

Therefore, by Bayes' Theorem,

$$\begin{aligned}\pi(\mu|t) &\propto \pi(\mu)f(t|\mu) \\ &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{n\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{n\tau}{2}(t - \mu)^2\right\} \\ &\propto \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}(t - \mu)^2\right\} \\ &\quad \vdots \\ &\propto \exp\left\{-\frac{D}{2}(\mu - B)^2\right\}\end{aligned}$$

where B and D are as in Equation (2.21), with t replacing $\bar{x}$; that is, $\mu|t \sim N(B, 1/D)$, the same distribution as $\mu|\mathbf{x}$.